HIV-1 Reverse Transcriptase: Polymerization Properties of the p51 Homodimer Compared to the p66/p51 Heterodimer†

M. R. Bavand,‡ R. Wagner,§ and T. J. Richmond*,‡

Institut für Molekularbiologie und Biophysik, ETH-Hönggerberg, CH-8093 Zürich, Switzerland, and Institut für medizinische Mikrobiologie und Hygiene, 8400 Regensburg, Germany

Received May 13, 1993; Revised Manuscript Received July 26, 1993®

ABSTRACT: The polymerase activity of the p51 homodimeric form of HIV reverse transcriptase was characterized by activity gel analysis, steady-state kinetic measurements, and processivity assays, and the activity was shown to be highly similar to that for the p66/p51 heterodimer. Recombinant 51- and 66-kDa reverse transcriptase proteins were individually expressed from an HIV-1 *Pol* gene having an accumulation of natural amino acid mutations compared to the BH10 clone (Ratner *et al.*, 1985). The preparation of an active p51 homodimer critically depended on low temperature during its expression in bacterial cultures. Activity gel analysis demonstrates that refolded p51 protein derived from denatured p66/p51 heterodimer yields an active polymerase. The p51 homodimer has approximately one-half the activity and processivity of the heterodimer, while both enzymes have similar thermostability. Steady-state measurements reveal no significant differences in apparent affinities for substrate or homopolymeric template–primer, suggesting that the subunits in both enzyme forms have similar conformations. Template challenge experiments show that the off-rates for template–primer are lower, but as indicated by primer extension analyses, processivity is less for p51 homodimer. These results show that the RNase H domain is not essential for the assembly of the functional polymerase, but suggest that it enhances processivity.

Reverse transcriptase (RT) from human immunodeficiency virus (HIV-1), a prime target for drugs which interfere with the onset of acquired immune deficiency syndrome (AIDS), has been the subject of intense research [for reviews, see Barber *et al.* (1990) and Jacobo-Molina and Arnold (1991)]. When RT is extracted from virions or infected cells and subjected to denaturing SDS gel electrophoresis, protein bands of MW 66 000 and 51 000 are detected in equal quantities, suggesting that a heterodimeric enzyme constitutes an active complex *in vivo* (DiMarzo Veronese *et al.*, 1986). The larger subunit possesses both an N-terminal polymerase and a C-terminal RNase H domain, while the smaller subunit contains only the former. This heterodimeric RT is thought to mature by a proteolytic processing of the *gag–pol* fusion protein, generating an asymmetric p66 homodimer from which only one p66 subunit is further cleaved by the viral protease (DiMarzo Veronese *et al.*, 1986).

Reverse transcription is a multistep process that includes polymerization, nuclease digestion, and template-jumping steps between two viral RNA molecules (Panganiban & Fiore, 1988). This process takes place mainly within the cytoplasm of infected cells (Weiss *et al.*, 1982), although some reverse transcription seems to be initiated before or during assembly of the virion (Lori *et al.*, 1992). The RT proteins performing these steps in reverse transcription and mutants thereof have been expressed in different heterologous systems using recombinant techniques and then characterized biochemically [Barber *et al.* (1990); Jacobo-Molina and Arnold (1991) and references therein]. The RT complex having the highest activity and stability in these studies is the p66/51 heterodimer, and therefore it seems to be the most relevant form *in vivo*.

The activity of HIV-1 RT proteins has been strictly correlated with the dimer state, and the association constant of the heterodimer has been reported to be approximately $10–10^3$ times higher than that for the p66 homodimer and somewhat larger yet with respect to the p51 homodimer (Becerra *et al.*, 1991; Restle *et al.*, 1990). Cross-linking studies with a substrate analog have shown that the active site is confined to the p66 subunit of the heterodimer (Chen *et al.*, 1991), and the crystal structure shows that the p51 subunit of that molecule is locked in an inactive configuration (Kohlstaedt *et al.*, 1992). The crystal structure has revealed that the subunit interactions are located mainly in the connection and thumb subdomains of the p51 subunit, in the connection and palm subdomains of the p66 subunit, and to a lesser extent, in the tip of the thumb subdomain of the p51 subunit and the RNase H region of the p66 subunit.

Since the dimeric state is a prerequisite for the polymerase function, the steady-state and pre-steady-state kinetics have been investigated only for the p66 homodimer and p66/51 heterodimer (Majumdar *et al.*, 1988; Huber *et al.*, 1989; Kedar *et al.*, 1990; Reardon & Miller, 1990; Krug & Berger, 1991; De Stefano *et al.*, 1991; Reardon *et al.*, 1991; Reardon, 1992). Little attention has been given to the characterization of the p51 homodimer, since expression has yielded mainly an inactive or weakly active protein (Hansen *et al.*, 1988; Hizi *et al.*, 1988; Mueller *et al.*, 1989; Sobol *et al.*, 1991). On the other hand, several reports have shown that p51 molecules may have polymerase activity (Restle *et al.*, 1990; Dirani-Diab *et al.*, 1992; Hostomsky *et al.*, 1992). Since the role of the p51 molecule engaged as either a homo- or heterodimer is unclear for RT function, we have studied the enzymatic activity of the p51 homodimer.

In this report, we describe the cloning, overexpression, and purification of recombinant p66 and p51 HIV-1 RT proteins and the preparation of an active p51 homodimer as well as a defined-sequence p66/51 enzyme. The results of activity gel and heat denaturation analyses and steady-state kinetic

‡ Support of this work was provided by Boehringer-Mannheim GmbH.
* Author to whom correspondence should be addressed.
‡ ETH-Hönggerberg.
§ Institut für medizinische Mikrobiologie und Hygiene.
® Abstract published in *Advance ACS Abstracts,* September 15, 1993.

0006-2960/93/0432-10543$04.00/0

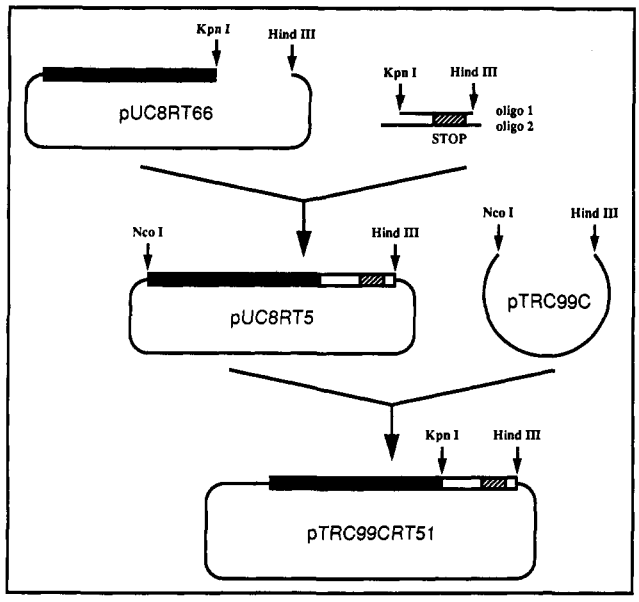

FIGURE 1: Construction scheme of the plasmid for the overproduction of p51. Construction of expression vector pTRC99CRT51 for the expression of recombinant p51RT. Detailed descriptions of constructions are given in the Experimental Procedures.

and processivity measurements for both enzymes demonstrate that the RNase H domain is not essential for the formation of an active enzyme. However, the RNase H domain may modify the polymerase function with template–primer since initiation is delayed and processivity is reduced somewhat for the homodimer compared to the heterodimer. Since the deduced amino acid sequence of the RT gene used here has a significant number of natural mutations, the kinetic parameters for p66/51 are compared to those for earlier RT clones.

EXPERIMENTAL PROCEDURES

Virus DNA Source and Plasmids. The p66 and p51 RT genes that were used in our studies were derived from HIV-1 strain WMF 1.13 proviral DNA, isolated from peripheral blood lymphocytes (Saag *et al.*, 1988), and cloned into a bacterial expression vector (pKK233–2RT5p66), allowing the expression of the p66 RT gene (Soutschek *et al.*, 1989). The p51 RT gene was derived from the p66 RT gene and was inserted into the bacterial expression vector pTRC99C (Pharmacia). The expression vector pTRC99CRT51 was constructed as follows: Vector pUC8RT66 (Soutschek *et al.*, 1989) was digested with *Kpn*I and *Hin*dIII and ligated in the presence of annealed DNA oligonucleotides 1 and 2 with the following sequences (Figure 1):

5′-CAGTTAGAGAAAGAACCCATAGTAGGA
GCAGAAACCTCCTAGA-3′ (oligo 1)

5′-AGCTTCTAGGAGGTTTCTGCTCCTACTA
TGGGTTCTTTCTCTAACTGTAC-3′ (oligo 2)

The resulting construct, termed pUC8RT5, was digested with *Nco*I and *Hin*dIII, releasing a 1300-bp fragment with the p51 gene that was inserted into an *Nco*I/*Hin*dIII-digested vector pTRC99C, yielding p51 expression vector pTRC-99CRT51. This construct encodes the same amino terminus as the above p66 gene, where a lysine has been substituted for a proline. The carboxy-terminal amino acid was chosen to be Phe 441, as this appears to be the major cleavage site *in vivo* (Graves *et al.*, 1990). The p66 and p51 RT genes were sequenced by dideoxy sequencing (Sanger *et al.*, 1987), and the deduced amino acid sequence was obtained using the sequence analysis software package GCG (Devereux *et al.*, 1984).

Expression of the HIV-1 RT Subunits. Escherichia coli strain HB101 was used as a host strain to express the p66 and p51 proteins encoded by pTRC99CRT51. Bacterial growth was performed in 2× TY medium (Sambrook *et al.*, 1989). Large-scale expression of both the p51 and p66 proteins was as follows: Competent HB101 cells were transformed with either expression plasmid pKK233–2RT5p66 or pTRC-99CRT51. A single transformant was used to inoculate a 10-mL culture of 2× TY medium containing 100 μg/mL ampicillin (2× TY/Amp), and the culture was grown at 30 °C until it reached an OD_{600} of 0.8. This culture was then transferred to a 2-L flask containing 500 mL of 2× TY/Amp and was grown overnight at 30 °C. Subsequently, 500 mL of culture was used as an inoculum for 24 2-L Erlenmeyer flasks containing 700 mL of 2× TY/Amp medium. Cultures were grown at 30 °C at a shaker speed of 200 rpm until they reached OD_{600} values of 1.4. The temperature was lowered to 20 °C, and expression was induced with 100 μg/mL IPTG for an additional 17 h. Cells were harvested by centrifugation and stored frozen at −70 °C. The yield was typically 100 g of wet biomass.

Purification of the p66 and p51 RT Subunits. Frozen cells (100 g wet weight) were homogenized in 250 mL of lysis buffer consisting of 50 mM Tris Cl (pH 8.0), 200 mM NaCl, 1 mM EDTA, 10% (v/v) glycerol, and 2 mM DTT and were combined with 50 mL of lysis buffer containing 380 mg of lysozyme (Pharmacia). The suspension was stirred at room temperature for 20 min and then for another 20 min at 4 °C. Thereafter, the suspension was made 1% (v/v) Nonidet P-40 and 1 mM in phenylmethanesulfonyl fluoride (PMSF) and was homogenized using an Ultra-Turrax mixer (Janke & Kunkel KG, Staufen, Germany). Nucleic acids were removed by the addition of a 10% (v/v) poly(ethylene imine) solution (pH 8), to a final concentration of 0.7% (v/v), and the homogenized slurry was placed on ice for 10 min. After centrifugation at 12 000 rpm (HFA12.500 rotor, Heraeus) for 30 min at 4 °C, the supernatant was stored on ice and the pellet was resuspended in 200 mL of lysis buffer. This washing step was repeated, and the supernatants were then combined and diluted with ice-cold D-buffer (50 mM Tris Cl (pH 7.5) at 0 °C, 2 mM EDTA (pH 8), 100 μg/mL benzamidine, and 2 mM DTT) to an equivalent of 130 mM NaCl, as judged by conductivity measurements using a Bio-Rad Model 1710 conductivity meter. A wet volume of 100 mL of freshly regenerated phosphocellulose (Whatman P11) was added, and the suspension was allowed to stir at 4 °C for 1 h. All subsequent steps were performed at 4 °C. The phosphocellulose suspension was then vacuum-filtered using a large (13 cm diameter) glass frit (Schott) and subsequently washed four times with 250 mL of D-buffer and twice with 250 mL of D-buffer containing 130 mM NaCl.

Batch elution from the phosphocellulose was done by three cycles of 100 mL of D-buffer containing 2 M NaCl. The eluate was brought to 100 mM Tris Cl (pH 7.5), 5 mM EDTA, and 50% (saturated) ammonium sulfate and was stirred manually to reduce shearing until all the ammonium sulfate was dissolved. The suspension was centrifuged at 15 000 rpm (HFA22.50 rotor) for 15 min, and the pellet was dissolved in a minimal volume of D-buffer containing 1 M NaCl. This solution was dialyzed for 3 h against 2 L of D-buffer and then overnight against 2 L of Q-buffer (20 mM diethanolamine/

taurine, pH 9) containing 100 μg/mL benzamidine, 1 mM EDTA (pH 8), and 2 mM DTT. The resulting precipitate from overnight dialysis was removed by centrifugation at 15 000 rpm (HFA22.50 rotor) for 15 min. The supernatant was loaded onto a 2.5 × 10 cm column containing Q-Sepharose (Pharmacia), preequilibrated in Q-buffer. The Q column was washed with 10 column vol of Q-buffer, and proteins were eluted by a linear 50–150 mM NaCl gradient in Q-buffer (total volume 500 mL). Fractions of 10 mL were collected and analyzed by SDS–PAGE minigel (Cambridge Electrophoresis Ltd., England). Fractions enriched in RT proteins were pooled, diluted with D-buffer to 90 mM NaCl, and passed over a column packed with 5 × 5 cm Poly-U Sepharose (Pharmacia). The column was washed with 5 column vol of D-buffer, and elution was carried out with 600 mL of a linear gradient of 50–300 mM NaCl in D-buffer. Fractions were analyzed as described above, and pooled fractions were then diluted to 70 mM salt and passed over a 3 × 2.5 cm Heparin Sepharose (Pharmacia) column in a cycling mode overnight. Elution was carried out with 2 mM PIPES/NaOH (pH 6.8), 2 M NaCl, and 0.1 mM EDTA. The eluate was further concentrated by ultrafiltration with several buffer changes against storage buffer (20 mM Hepes/NaOH (pH 7.0), 100 mM NaCl, and 10% glycerol) in an Ultrathimble (Sartorius). The concentrated protein was brought to 50% (v/v) glycerol and was stored at −20 °C. Protein concentrations were determined by the Bradford assay (Bradford, 1976), using bovine serum albumin as a standard, or by using the extinction coefficient of p51 derived from amino acid analysis (see below).

Reconstitution of the p66/51 Defined-Sequence Heterodimer. The mixing of harvested cells, and subsequent normal purification of the separately expressed RT subunits, yields heterodimer directly. Each 100 g of bacterial cells (wet weight) containing separately expressed p66 and p51 subunits was mixed, and the cells were lysed as described above. Purification steps were essentially the same as for the isolation of homo- and heterodimer RT. Based on densitometric analyses of Coomassie-stained minigels (Figure 3), the p51 homodimer and p66/51 forms of HIV RT purified as described above were >98% pure.

Storage of Purified RT. After purification and concentration, RT was stored in 10 mM Hepes/NaOH (pH 7), 50 mM KCl, 50% (v/v) glycerol, and 5 mM DTT at −20 °C. All polymerization assays (see below) were performed either with defined-sequence p66/51 heterodimer obtained by separate expression of subunits and reconstitution or with defined-sequence p51 homodimer.

Activity Gel Analysis. The DNA-dependent DNA polymerase activities of the p51 and p66 homodimers and the p66/51 heterodimer were analyzed on activity gels. Activated calf thymus DNA (1140 units, Pharmacia) was included as template–primer in a 7 × 9 × 0.1 cm SDS gel containing 8% polyacrylamide, 0.5% bisacrylamide, and 0.1% SDS. A quantity of 30 μg of p51, p66, or p66/51 form of HIV RT was solubilized in loading buffer consisting of 4% (w/v) SDS, 0.125 M Bis-Tris (pH 6.8), 5%(v/v) β-mercaptoethanol, and 30%(v/v) glycerol and loaded without heating. The separation, subsequent renaturation, and assay were performed as described previously (Bavand *et al.*, 1989).

Amino Acid Analysis. Purified p51 protein (10 nmol) was hydrolyzed *in vacuo* at 110 °C for 24 h in constant-boiling HCl. Dried samples were dissolved in water containing norleucine as a standard, and analysis was performed with a Biotronic LC 3000 amino acid analyzer. The extinction coefficient of the p51 homodimer was calculated from the absolute amount of protein as ϵ_{p51} = 142315 L/mol·cm, consistent with a theoretical value of ϵ_{p51} = 124 420 calculated according to Gill and v.Hippel (1989). The amino acid analysis confirmed the deduced amino-terminal sequence of the p51 protein, starting with the sequence M-K-I-S-P.

Steady-State Polymerization Assays. Reactions were carried out in 50- or 100-μL volumes. The concentrations of recombinant HIV reverse transcriptase (p66/51 and p51 forms) and template–primer varied in the different experiments and are specified in the figure legends. The standard RT assay buffer contained 50 mM Tris Cl (pH 8.0), 50 mM KCl, 6 mM $MgCl_2$, 5 mM DTT, and 0.02 μg/μL bovine serum albumin (BSA), to which 35 μM dTTP and 100 nM poly(rA)oligo$(dT)_{12-18}$ were added. Poly(rA) and oligo-$(dT)_{12-18}$ were obtained from Pharmacia. The average base length of the homopolymeric poly(rA) template varied batchwise between 400 and 1100 nucleotides and was calculated from the sedimentation velocity ($S_{20,w}$) number of the product specification with the formula $\log M = 2.179(\log S_{20,w} + 1.523)$ (Eisenberg & Felsenfeld 1967). BSA was used to stabilize diluted enzyme in enzyme dilution buffer and was also present in the polymerization mix at the indicated concentration to prevent adsorption of enzyme to the reaction vessel at high dilution. Enzyme dilutions in RT buffer were made on ice and used immediately for the polymerization assays. Standard assays were for 10 min at 30 °C unless otherwise indicated. Data points represent the averages of 2–5 independent measurements. Reactions were carried out for the indicated times and stopped by the addition of 10 μL of 0.5 M EDTA. After the addition of 500 μL of 10% (w/v) trichloroacetic acid/2% (w/v) sodium pyrophosphate (TCA/PP_i), the acid precipitate was collected on Whatman GF/B filterdisks and washed with twice with 10 mL of ice-cold TCA/PP_i and twice with 3 mL of ice-cold absolute ethanol. Air-dried disks were counted using Hisafe II scintillation fluid (LKB, Bromma). The quenching factor of the glass filterdisks was taken into account; the counting efficiency was 65%. One unit is defined as 1 nmol of dTMP incorporation into poly(rA)oligo$(dT)_{12-18}$ at 30 °C for 30 min. At a concentration of 5×10^{-8} M, purified recombinant p51 had a specific activity of approximately 3000 units/mg, and the p66/51 defined heterodimer had a specific activity of approximately 6000 units/mg, which are comparable to the specific activities obtained by others (Restle *et al.*, 1990).

The dependency of the product formation on dTTP substrate concentration was measured in 100-μL assays containing 10 nM enzyme and 200 nM poly(rA)oligo$(dT)_{12-18}$ in RT assay buffer. Enzyme and template–primer were mixed in a 22-μL volume, and reactions were started by the addition of varying concentrations of Mg-dTTP in 78 μL of RT assay buffer. Reactions were stopped after 3 min at 37 °C, and product formation was assayed. Substrate initial velocity data were plotted v over v/s, and the K_m data were obtained from the linear regression curve.

For K_m determinations of template–primer, polymerization reactions containing 50 μL of 20 nM enzyme and 300 μM dTTP in 2× concentrated RT buffer were initiated by mixing with 50 μL of 2× concentrated poly(rA)oligo$(dT)_{12-18}$ solution at different concentrations. Reactions were stopped after 10 min at 30 °C, and product formation was assayed. K_m values were derived from Eadie–Hofstee plots.

Processivity Assays. Polymerase processivity was measured from primer extension reactions. A DNA oligo$(dT)_{14}$ was labeled with [^{32}P]ATP (Amersham) according to standard procedures (Sambrook *et al.*, 1989). End-labeled oligo$(dT)_{14}$

primer was mixed with an excess of unlabeled primer and hybridized to poly(rA) template in a 1:1 molar ratio in 5 mM Hepes/NaOH (pH 7.0) and 0.5 mM EDTA. The oligo(dT)$_{14}$ primer was made on an Applied Biosystems 380B DNA synthesizer and checked for purity by denaturing gel electrophoresis. Hybridization was performed by heating to 90 °C for 1 min and then slowly cooling to room temperature over a period of 20 min. Final concentrations in primer extension reactions were 500 nM template–primer, 30 nM purified RT, and 100 μM dTTP in standard RT assay buffer. Solutions were equilibrated at 30 °C for 20 s before the reaction was started by adding 2.2 μL of Mg-dTTP mix to 14.4 μL of a solution containing labeled template–primer and enzyme in RT assay buffer. Polymerization was stopped by adding 5 μL of formamide loading solution consisting of 98% (v/v) formamide, 10 mM EDTA, and 1 mg/mL each of xylenecyanol FF and bromophenol. The mix was heated for 1 min at 90 °C, and an aliquot of 6 μL was electrophoresed on a 12% polyacrylamide/7 M urea gel at 41 W. Exposure of gels to X-ray film was at −70 °C in the presence of an intensifying screen.

Processivity experiments under challenge assay conditions were similar to those described by Bryant *et al.* (1983), except that activated calf thymus DNA (Pharmacia) was used as a trap. A solution was prepared containing 30 nM poly(rA)-oligo(dT)$_{12-18}$ and 30 nM purified p51 or defined p66/51 RT in standard RT assay buffer and incubated for 5 min on ice and then for 2 min at 30 °C. Reactions were started by adding 16 μl of a solution containing 25 mM $MgCl_2$, 315 μM dTTP, and 2.5 μg/μL activated DNA to 34 μL of the above solution and were incubated at room temperature. Reactions were quenched at different times by adding 10 μLof 0.5 M Na EDTA (pH 8.0) and 500 μL of TCA/PPi solution. The acid precipitate was collected and washed on glass filters as described above. Control experiments consisted of either reactions where preincubations of enzyme and template–primer were carried out in the presence of a DNA trap or reactions without a DNA trap. Kinetic constants were derived graphically from Eadie–Hofstee plots and interactive curve-fitting using the data analysis and graphics program, Kaleidagraph (Synergy Software, Reading, PA).

RESULTS

Primary Sequence. The deduced amino acid sequence of the p66 RT gene of the WF1.13 isolate revealed 27 amino acid differences, listed in Table I, from that of the standard BH10 isolate (Ratner *et al.*, 1989). When compared to the amino acid sequences from 14 other HIV-1 isolates deposited in the EMBL database (not shown), 11 of these were found to be novel. When the mutations were assigned to individual structural subdomains (Kohlstaedt *et al.*, 1992), two mutational hotspots were apparent (Figure 2). One of these with nine mutations, is in the connection subdomain, which forms part of the dimerization interface of heterodimeric RT, and suggests that the precise orientation of the RNase H domain with respect to the p51 portion of a subunit is not critical to function. The second hotspot, with five mutations, is in the palm region near the thumb and forms part of the apparent template–primer binding cleft of the p66 subunit. Although this region is distant from the RNase H domain, a recombinant double mutant in which two highly conserved proline residues in the same region (Pro 225 and 226) were converted to threonine residues had little RNase H activity while they retained full polymerase activity (Hizi & Shaharabany, 1992). In this case, the RNase H activity appears to depend on RT or template–primer structure external to the RNase H domain.

Table I: Comparison of Amino Acid Differences Found between the RT Gene of the WF1.13 Isolate Used in This Study and the BH10 Isolate (Ratner *et al.*, 1985)[a]

position	WF1.13	BH10	position	WF1.13	BH10
1	K	P	334	H	Q
9	S	P*	338	S	T*
88	L	W*	365	E	V*
99	W	G*	370	K	E*
122	K	E	376	S′	T
155	V	G*	390	R	K
177	G	D	400	A	T
191	L	S*	403	I	T
207	A	Q	460	D	N
211	K	R	468	S	P*
214	F	L	471	D	N
245	M	V	483	H	Y
293	V	I	556	T	I*
333	E	G*			

[a] Amino acid differences in the primary sequence of the two isolates are shown in a one-letter code together with the position. Novel amino acid mutations are indicated with an asterisk (*).

Expression and Purification of the RT Forms. Several attempts to construct a p51 RT gene in pKK233 were unsuccessful because of plasmid loss. Using pTRC99C, a more stringently regulated pKK plasmid, in order to avoid the production of p51 during cell growth, we succeeded in making large quantities of p51. Even with the use of this more tightly regulated plasmid, the overexpression of soluble p51 protein critically depends on both growth and induction of bacterial cultures at temperatures at or below 30 °C. Growth at higher temperatures leads to loss of the expression plasmid (ampicillin concentration in the medium maintained at 100 μg/mL). Growth at 30 °C to the late logarithmic phase (OD_{600} = 1.4) and induction at 20 °C lead to the accumulation of optimal amounts of soluble, recombinant p51 protein, although insoluble p51 is always present as well. Induction at lower temperatures does not improve the yield (data not shown). On the basis of densitometric scans of protein gels of induced and uninduced cultures, we estimate that 15% of the total soluble cell protein is expressed as p51 in induced cultures.

The purification steps for recombinant p66, p51, and p66/51 proteins were essentially identical. For the preparation of a defined heterodimeric p66/51 RT, equal amounts of cell pellets containing individually expressed proteins were mixed and processed together. High yields of purified homodimeric and heterodimeric RT proteins were obtained when cell pellets were processed directly after harvesting. The purification protocol reproducibly yielded about 100 mg of purified protein from about 100 g of wet cells or 18 L of culture media. The individual purification steps were monitored by SDS–PAGE (Figure 3C).

Although approximately 50% of p51 RT aggregated irreversibly during overnight dialysis against Q-buffer, several contaminants were coprecipitated to yield p51 protein which was about 95% pure. The solubilization of the ammonium sulfate precipitate was facilitated by a buffer solution containing 1 M NaCl, and subsequently, after dialysis to low salt concentration, about 20% of the p51 was found in the Q-Sepharose column flow-through fraction. This material showed 50-fold lower specific activity than the eluate, and therefore it may consist mainly of aggregated p51 protein. We estimate that from the total p51 species expressed under optimized conditions, only about 20% remained as an active fraction that bound to the Q-Sepharose resin. The dialysis step and the Q-Sepharose step are important in selecting a conformationally homogeneous p51 enzyme.

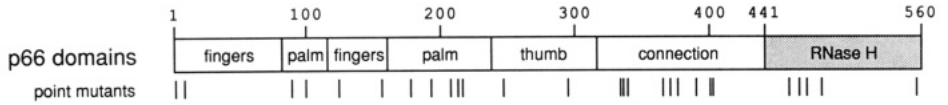

FIGURE 2: Polymerase and RNase H domains of RT. The locations of subdomains of the primary structure of p66 HIV-1 reverse transcriptase with respect to amino acid number are shown and named according to Kohlstaedt *et al.* (1992). The approximate locations of amino acid mutations of the RT found in the WF1.1.3 clone used in this study are indicated (|). The p51 protein terminates at amino acid 441.

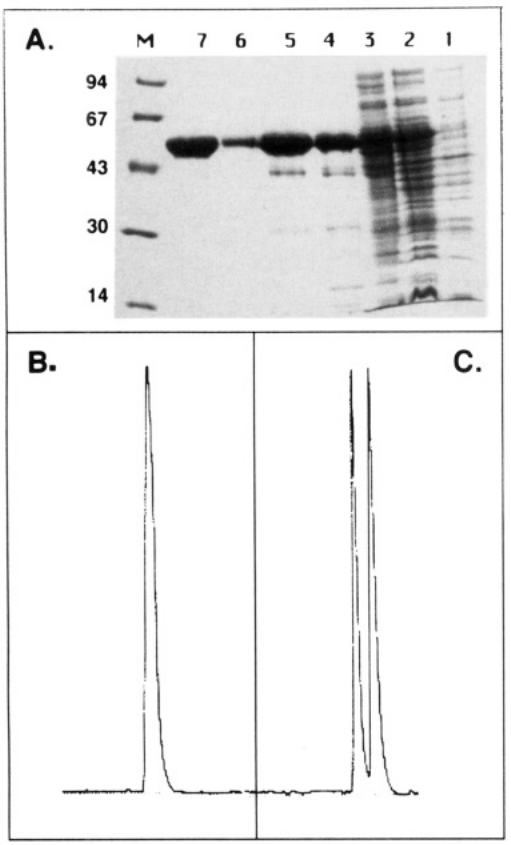

FIGURE 3: Purity of the recombinant RT. (A) SDS–PAGE of recombinant p51 during its purification: lane 1, uninduced culture; lane 2, induced culture (17 h, 20 °C); lane 3, lysis, supernatant; lane 4, phosphocellulose elution; lane 5, Q-Sepharose loading; lane 6, Q-Sepharose elution/Poly-U Sepharose loading; lane 7, Poly-U Sepharose elution; M, marker proteins. (B) Densitometric scan of purified p51 RT. (C) Densitometric scan of reconstituted and purified defined-sequence p66/51 RT.

After Poly-U Sepharose fractionation, the RT proteins were greater than 98% pure as evaluated from densitometric scans of Coomassie-stained SDS gels with p51 (Figure 3A) and defined p66/51 (Figure 3B). Furthermore, the scans of the purified heterodimer (Figure 3C) show that the reconstituted heterodimer contains an equimolar ratio of 66- and 51-kDa subunits. Samples of purified homodimeric p51 which were passed over columns of Superose 6 and 12 connected in series gave a single symmetrical peak eluting slightly earlier than that for a comparable run of the p66/51 (not shown). Since we obtained a highly active p51 protein that was homogeneous in gel filtration experiments, we assumed that our purification protocol also selects for homogeneous homodimer. The purified p51 homodimer maintained over 90% of its catalytic activity after 3 weeks of storage.

Activity Gel Analysis of p51, p66, and p66/51 RT Forms. Activity gel analysis relies on denaturation of a protein without heat and subsequent SDS gel electrophoresis, followed by renaturation and assay of its activity in the gel. There have been conflicting results regarding the activity of p51 with this assay system (Hansen *et al.*, 1988; Starnes *et al.*, 1988; Hu & Kang, 1991). In one case, the p51 was either inactive or its activity was lost during the unfolding/refolding steps (Hansen *et al.*, 1988). We repeated this analysis and found that, independent of the starting form of the RT, the p66 and p51 subunits renatured in the gel both show activity, with the p66 RT having somewhat higher activity as visualized on the autoradiogram (Figure 4). Quantification of band intensities from autoradiograms indicates that the p66 form is 1.5 times more active than the p51 form. Only the bands shown with the mobilities of monomer species appeared on the gel. This experiment confirms earlier data for p51 (Hu & Kang, 1991) and demonstrates that homodimeric and heterodimeric RT forms can be denatured in SDS and renatured to yield active homomeric enzymes. The inactive form of p51 in the heterodimer can be converted to an active form in the p51 homodimer.

Characterization of the Polymerization Reaction for p51 and p66/51. In optimizing the reaction conditions for a polymerization system, we have analyzed the effect of KCl on the activity of both enzyme forms. An increase in KCl concentration inhibits p51 activity, particularly at concentrations above 40 mM (Figure 5a). At 100–120 mM KCl, the concentration used in most RT assays, p51 has only about 40% of its maximal activity. In contrast, p66/51 was most active at 120 mM KCl. At approximately 50 mM KCl, both RT forms were equally active in polymerization. Dependence on Mg^{2+} ion concentration was the same for both enzyme forms with an optimum at 6 mM (not shown).

When product formation over time was assayed (Figure 5b), both p51 homodimer and p66/51 heterodimer showed a linear course of product formation for the first 10 min, with average velocities of 26 ± 3.7 and 31 ± 3.9 pmol of product/min, respectively, with p51 showing a delayed onset of polymerization. At longer reaction times (>10 min), the rate of product formation was constant for the p66/51 form, whereas p51 continued with a diminished but linear rate.

Extrapolation of the product formation curve of the steady-state reaction to zero time shows a nonzero intercept on the product axis for p51 (Figure 5b, time 10–30 min). This may be interpreted as the participation of more than one kinetic event in the observed rate for p51 or a difference in the stabilities of both enzyme forms. Both enzyme forms revealed a similar decrease in their polymerization activity when preexposed to heat (30 °C) in the presence of template–primer (Figure 5c), however, which indicates that the lack of the RNase H domain in the p51 homodimer does not accentuate the heat destabilization of the polymerase domain in the presence of template–primer.

The substrate dependence on product formation for the p51 and p66/51 enzymes was also examined (Figure 5d). The reaction was started by Mg-dTTP of concentrations of 10–240 μM added to a solution containing premixed enzyme and template–primer. Product formations for reaction times of 4 and 30 min were recorded. Both p51 and p66/51 enzymes show similar product formation until approximately 40 μM dTTP, where inhibition is then observed for both enzyme forms, consistent with earlier results (Furman *et al.*, 1991). However, a further increase in the substrate concentration leads to the apparent alleviation of substrate inhibition, as product for-

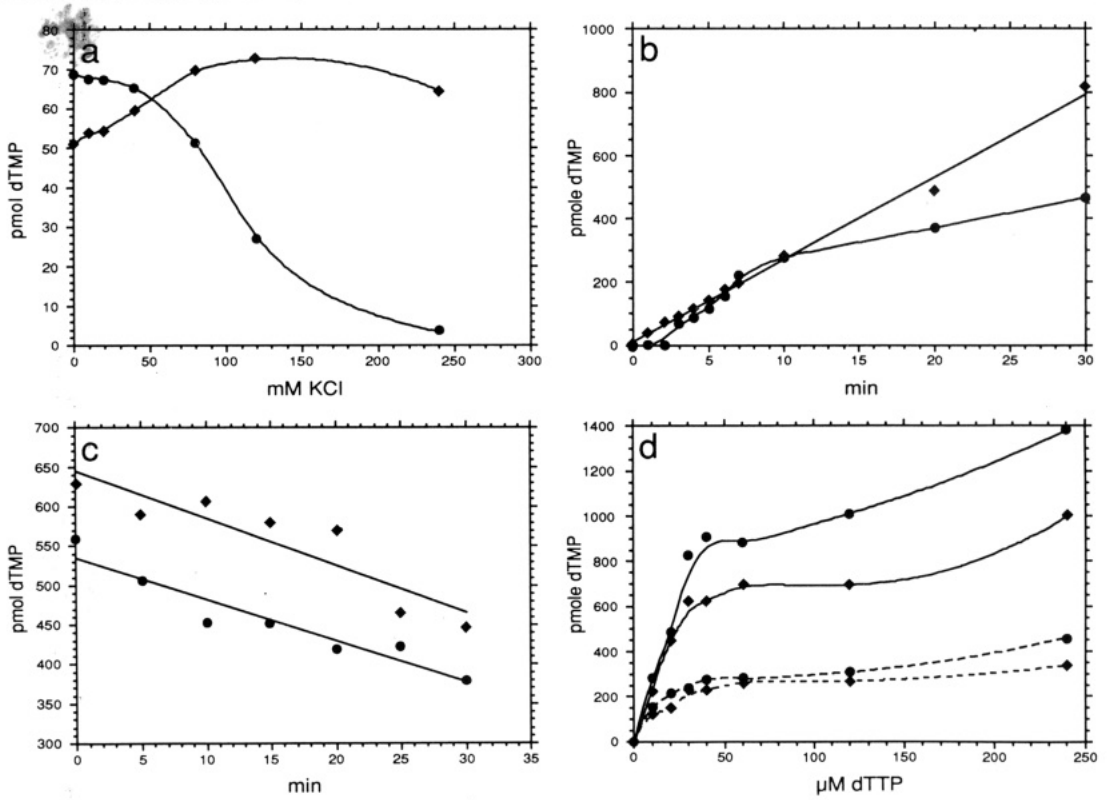

FIGURE 4: Characterization of the p51 homodimer and p66/51. Both the p51 homodimer (●) and p66/51 heterodimer (◆) were assayed for product formation. (a) Effect of KCl on polymerization activity. Reactions containing 100 μM dTTP, 100 nM poly(rA)oligo(dT)$_{12-18}$, 20 nM enzyme, and 0–240 mM KCl were incubated at 30 °C for 10 min. (b) Time course of poly(rA)-directed poly(dT) synthesis. Reactions containing 200 nM poly(rA)oligo(dT)$_{12-18}$, 150 μM dTTP, and 30 nM enzyme were stopped at the indicated time points. (c) Heat-inactivation profile. Mixtures containing 30 nM enzyme and 200 nM poly(rA)oligo(dT)$_{12-18}$ (no Mg^{2+}) were held at 30 °C for the indicated times prior to intiating the reaction by adding Mg-dTTP to 150 μM. (d) Substrate inhibition. Reactions containing 100 nM poly(rA)oligo(dT)$_{12-18}$, 30 nM enzyme, and 0–240 μM dTTP substrate were allowed to proceed at 30 °C for 4 (- - -) and 30 min (—).

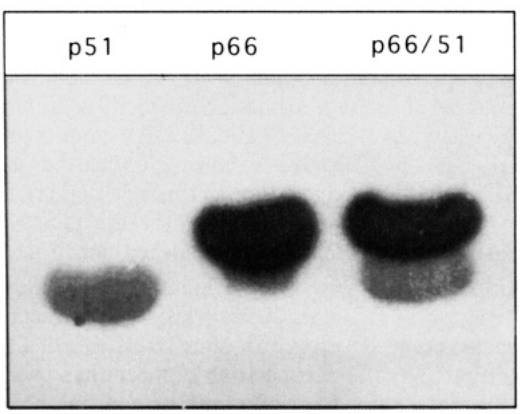

FIGURE 5: Activity gel analysis of the three RT forms. The autoradiogram shows the three RT forms as indicated. After separation, the proteins were renatured in the gel and the polymerase activity was assayed. Details are described in the Experimental Procedures.

mation again increases. The change in the rate of product formation at approximately 50 μM substrate suggests that there may be at least two modes of substrate interaction for these enzymes.

Apparent Michaelis–Menten kinetics for the polymerase reactions of both p51 and p66/51 RT were obtained for poly(rA)oligo(dT)$_{12-18}$ and dTTP as reactants. The dTTP substrate concentration was varied between 0 and 40 μM for short reaction times, and template–primer was held at a saturation concentration of 200 nM throughout (data not shown). The K_m values for dTTP as substrate derived from Eadie–Hofstee plots are 6.5 ± 0.6 μM for p51 and 4.6 ± 0.1 μM for p66/51. The k_{cat} values are 0.78 ± 0.03 s^{-1} for p51 and 1.03 ± 0.02 s^{-1} for p66/51, consistent with K_m and k_{cat} values previously measured for the heterodimer (Reardon & Miller, 1990; Furman *et al.*, 1991). The values of k_{cat}/K_m for the p51 and p66/51 enzymes were 0.12×10^6 and 0.22×10^6 s^{-1} M^{-1}, respectively, and correlate well with the approximately 2-fold higher specific activity of p66/51.

Since both the polymerase and RNase H domains combine to form the template-binding cleft in the heterodimer (De Vico *et al.*, 1991; Kohlstaedt *et al.*, 1992), we asked whether the reduced activity of p51 homodimer was due to altered polynucleotide binding. Steady-state kinetic measurements were made for both p66/51 and p51 RT varying the template–primer concentration, and values for kinetic constants were derived from Eadie–Hofstee plots. K_m values of 32.7 ± 1.8 nM for p51 and 20.4 ± 3.8 nM for p66/51 were obtained for template–primer at 150 μM dTTP. Values from 3.4 to 59 nM for the heterodimer and poly(rA)oligo(dT)$_{12-18}$ have been described (Huber *et al.*, 1988; Dudding & Mizrahi, 1991; Reardon & Miller 1990), indicating that subtle changes in reaction conditions may have a profound influence on template–primer binding.

The kinetic results obtained here indicate that, aside from subtle differences, the interactions of both enzyme forms with substrate and polynucleotide template–primer must be highly similar and that the polymerization function can be preserved in the p51 homodimer.

Characterization of the Processivity of p51 and p66/51. The processivity of RT, that is, the number of nucleotides polymerized during a single binding event, was evaluated by enzyme trapping in challenge assays and direct product analysis of extended primers. The distributions of DNA fragments, synthesized by both RT forms in a short time interval using

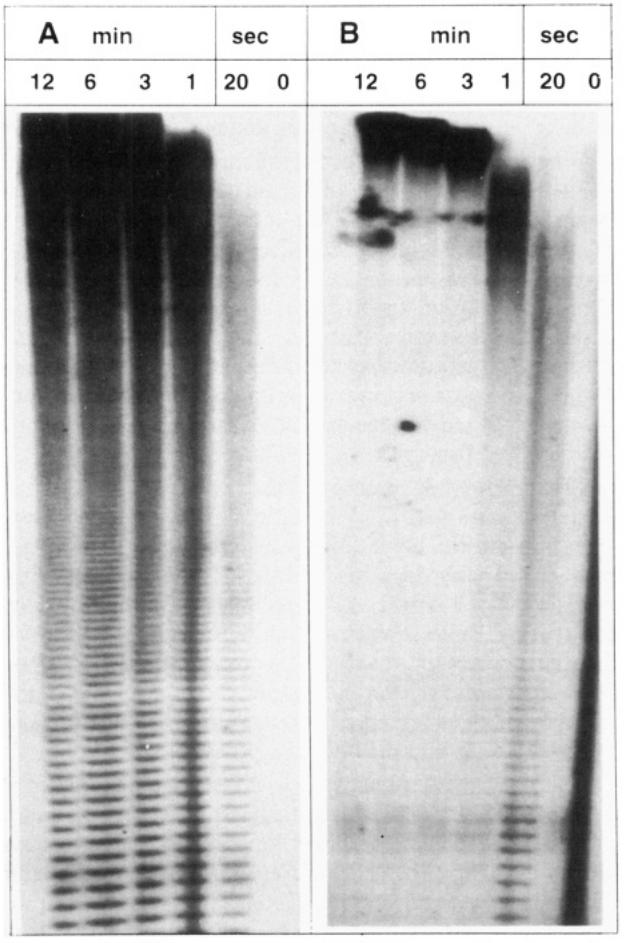

FIGURE 6: Time course of DNA synthesis using poly(rA)oligo([^{32}P]dT)$_{14}$: (A) p51 homodimer; (B) p66/51 heterodimer. Reactions containing 500 nM annealed hybrid ^{32}P-labeled oligo(dT)$_{14}$poly(rA) (1:1 molar ratio), 30 nM enzyme, and 100 μM dTTP in RT buffer were stopped at the indicated times, and the products were separated on 12% polyacrylamide/7 M urea gels as described in the Experimental Procedures. The average length of the templates used was approximately 400 nucleotides. The products with the highest mobility in each lane are 19 ± 2 nucleotides in length.

^{32}P-labeled oligo(dT)$_{14}$ heat annealed to poly(rA) and extended by addition of dTTP, are qualitatively the same (Figure 6, 20-s lanes). During the first 20 s, the maximum length of extension is approximately 230 nucleotides, as determined from a plot of elongated primer length versus primer migration distance on denaturing PAGE gels (not shown). This value corresponds to a rate of incorporation of 10–13/s, consistent with previous results for the heterodimer (Huber *et al.*, 1989). The value obtained from this experiment, from which only the maximally extended product is considered, is approximately 10-fold larger than the result from steady-state kinetic measurements, which measure the average value over all extension products. The distribution of bands appearing as a ladder indicates that the enzymes frequently dissociate from the template and that the number of reinitiations is too infrequent to complete the extension of all of the primer molecules in the presence of excess template (500 μM).

At longer times the rate of reinitiation apparently affects the appearance of the product distribution patterns (Figure 6A,B, 1–12 min). After 1–3 min of synthesis, the extended primers approach the average maximum size for the template of approximately 400 nucleotides (see Experimental Procedures), as determined by extrapolation from primer length versus migration plots. However, the p51 reaction products

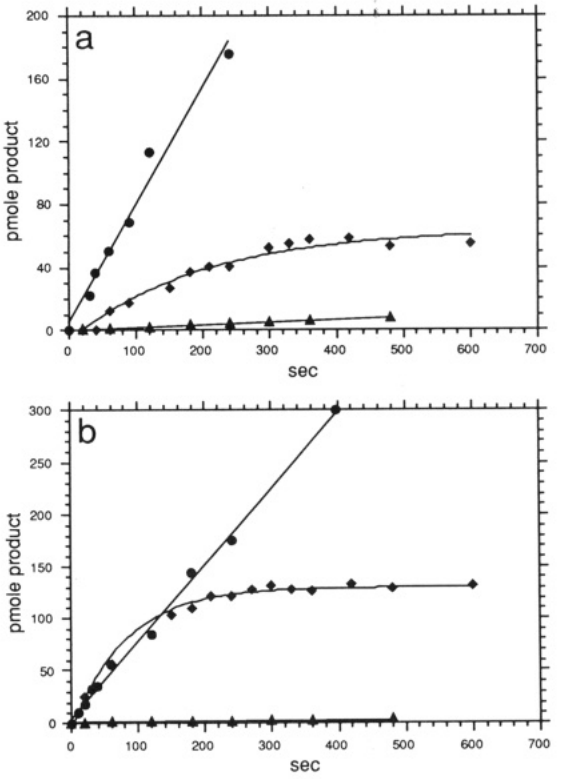

FIGURE 7: Time course of DNA synthesis from preformed p51 RT–poly(rA)oligo(dT)$_{12-18}$–enzyme complexes: (a) p51 homodimer; (b) p66/51 heterodimer. Reactions containing 30 nM enzyme and poly(rA)oligo(dT)$_{12-18}$ were carried out as described in the Experimental Procedures. ◆: Template–primer and challenger DNA were added simultaneously at the start of the reaction. ●: No challenger DNA were added. ▲: Challenger DNA and template–primer were preincubated together before the start of the reaction.

continue to show a ladder formation of elongated primer at all time points (Figure 6A, 1–12 min), whereas elongation by the heterodimer gave mainly full-length products (Figure 6B, 1–12 min). The primer length dependence of elongation noted for p66/51 by Reardon *et al.* (1991) is evident as some short products (≤20-mer) remain even at the longer times.

A template challenge assay (enzyme-trapping) was used to compare the processive properties of the two RT forms. Activated calf thymus DNA was added to the template–primer–enzyme complex to act as the trapping agent, *i.e.*, to provide an excess of 3′-OH groups for polymerase binding. The conversion of labeled primer to an extended, acid-insoluble form was followed. The rate of polymerization in the challenged reaction was considerably slower for p51 than for p66/51 RT and slower overall than that of the unchallenged reaction (Figure 7a,b). For both RT enzymes, a single cycle of processive synthesis required several minutes, consistent with results obtained for the p66/51 heterodimer (Huber *et al.*, 1989). A delay of approximately 30 s in the onset of the polymerization reaction after the addition of substrate is observed for p51 (Figure 7a). At the plateau levels after all enzyme is trapped, 55 and 123 pmol of dTMP were incorporated for p51 (after 6 min) and p66 (after 4 min), respectively. The background incorporation due to the trap DNA is approximately 9% and 3% for p51 and p66/51, respectively, as determined in a separate experiment in which trap DNA was added before the reaction was initiated. In the absence of challenger DNA, the reactions show constant polymerization rates for several minutes and reach a level of incorporation of more than 270 and 180 pmol of dTMP for p51 and p66/51, respectively. The ratio of the incorporation measured in the unchallenged to the challenge assay gives an

estimate of the number of cycles of initiation per template–primer under these conditions: 5 for p51 and 1.5 for p66.

The processivity of a polymerase is defined as the number of nucleotides polymerized during a single binding event or cycle and describes the competition between elongation in the bound state and dissociation from template–primer. In a challenge assay, where free enzyme is trapped efficiently, the following approximate equation describes the time dependence of product formation:

$$\frac{N(t)}{[ED]_0} = \frac{k_{cat}}{k_{off}}(1 - e^{-k_{off}t}) \quad (1)$$

where $N(t)/[ED]_0$ is the average number of dNMP incorporated into a DNA polymer per initial enzyme–DNA complex as a function of time (Bryant *et al.*, 1983). At long reaction times ($t \approx \infty$), this relationship yields a limiting value for the incorporation of substrate dTTP which can be evaluated from the plateau level for product formation (Figure 7a,b). This limiting value for $N/[ED]_0 = k_{cat}/k_{off}$ is defined as the processivity. Values for k_{cat} and k_{off} were derived by curve fitting the product formation data to eq 1. In order to allow for delayed onset of polymerization, data obtained for the p51 enzyme were fit to eq 2 by iterative substitution of k_{cat}, k_{off}, and c to obtain an optimized curve fit, where c is the time interval for the delayed onset of polymerization.

$$\frac{N(t)}{[ED]_0} = \frac{k_{cat}}{k_{off}}(1 - e^{-k_{off}(t-c)}) \quad (2)$$

A value for $N/[ED]_0$ of 43 was obtained for p51 (c = 19 s) and of 82 for p66/51, showing that the p51 enzyme is about one-half as processive as the heterodimeric enzyme, which is in good agreement with earlier processivity values for the heterodimer of 80–100 nucleotides for homo- and heteropolymeric templates (Reardon *et al.*, 1991; Huber *et al.*, 1989). The average processivity of p51 is similar to that of the Klenow fragment of DNA polymerase I, which has been shown to be 40–50 nucleotides (Bryant *et al.*, 1983). The polymerization reaction of p51, but not p66/51, was inhibited under challenge assay conditions, as seen by the lower rate of synthesis in the challenge versus the unchallenged assay. Therefore, the processivity value obtained for p51 may represent a lower limit.

The catalytic and dissociation rate constants for the enzyme–polymer complex were determined by fitting eq 2 to the product formation data (Figure 7), yielding values $k_{cat} = 0.46 \pm 0.01$ s^{-1} and $k_{off} = 0.007 \pm 0.001$ s^{-1} for the p51 homodimer and $k_{cat} = 1.68 \pm 0.03$ s^{-1} and $k_{off} = 0.012 \pm 0.001$ s^{-1} for the p66/p51 heterodimer. Therefore, the catalytic turnover of the heterodimer is 3-fold faster than that of the p51 homodimer, and it dissociates from the elongation complex approximately twice as fast. When this assay was performed with p51 enzyme that had been stored for 3 months at –20 °C, the kinetic parameters indicated that the enzyme had become less active and less processive ($k_{cat} = 0.127$ s^{-1}, $k_{off} = 0.009$ s^{-1}, $N/[ED]_0 = 9$; data not shown).

DISCUSSION

Expression and characterization of the p66 homodimeric and p66/51 heterodimeric RTs have been described by several groups, but few reports exist that deal with the characterization of p51. Some studies found the p51 protein to be poorly active, which was accounted for by the predominance of inactive monomers in the preparations (LeGrice *et al.*, 1991; Hansen *et al.*, 1988; Hizi *et al.*, 1988; Starnes *et al.*, 1988; Tisdale *et al.*, 1988; Mueller *et al.*, 1989). On the other hand, other reports describe polymerase activity for p51 when assayed on activity gels (Hu & Kang, 1991) or in solution (Dirani-Diab *et al.*, 1992; Restle *et al.*, 1990, Hostomsky *et al.*, 1992), although no details of expression and purification were given.

When we initially attempted to express p51 at temperatures of 30–37 °C, a product was obtained which was mostly inactive or insoluble. By lowering the induction temperature to 20 °C, we obtained an enriched fraction of soluble, highly active product. Apparently, the extent to which active p51 is expressed relative to other inactive p51 species critically depends on the temperature during growth and induction of the bacterial culture. As noted earlier (Schein & Noteborn, 1988), temperatures lower than that optimal for growth may be generally advantageous for expression and folding of recombinant proteins in *E. coli*.

The purification procedure used here selects for water-soluble species with high specific activity. The active p51 species is most likely a dimer, since the dimeric state of the protein and polymerase activity are strictly correlated (Restle *et al.*, 1990). Purified p51 homodimer maintains its activity for at least 3 weeks when stored in 50% glycerol at –20 °C at a concentration above 60 mg/mL. Inconsistent results reported previously for the characterization of p51 activity may be explained by expression conditions leading to different ratios of active and inactive p51 proteins.

When recombinant bacteria with separately expressed p66 and p51 subunits were mixed and lysed, and the proteins were processed together, heterodimer formation occurred consistent with previous results (Howard *et al.*, 1991; Le Grice *et al.*, 1991; Becerra *et al.*, 1991). Heterodimer was slow to form (15% full activity after 3 h at 4 °C) when separately purified active p51 and p66 preparations were combined, suggesting that p51 and/or p66 homodimers do not rapidly dissociate and reassociate as heterodimer. The formation of a functional heterodimer from purified monomeric subunits is associated with significant conformational changes in the polypeptide backbones of both p66 and p51 monomers and requires a significant incubation time to reach maximal activity (Anderson & Coleman, 1992).

The activity gel analysis (Figure 4) shows that the p51 subunit can refold and assemble as an active polymerase from denaturating conditions without the aid of the p66 subunit. It is also clear that a p66 subunit with mutations in essential amino acids of the polymerase active site cannot be compensated for by p51 in forming an active heterodimeric RT (LeGrice *et al.*, 1991). These results suggest that the p66 subunit induces or selects the inactive or locked conformation of the p51 subunit for assembly of the heterodimeric RT. Directly or indirectly, this must be due to the presence of the RNase H domain in p66. On the other hand, we show that p51 alone can assemble into a functional polymerase with enzymatic properties similar to those of the heterodimer. These properties include heat stability, k_{cat}, K_m, and processivity. Therefore, the RNase H domain is not essential for the formation of an active polymerase.

The RNase H region contacts the thumb subdomain of the p51 subunit in the heterodimer, and it appears that these interactions between the two domains are substantial [see Figure 3 of Kohlstaedt *et al.* (1992)]. Approximately 32% of the accessible surface area of the individual subunits buried in the dimerization interface of the p66/51 X-ray structure is due to the RNase H domain (T. A. Steitz, personal communicaton). These contacts could well account for the large difference between the dimer stability of the p51

homodimer and the heterodimer observed in several studies [*e.g.*, Restle *et al.* (1990)].

The lack of the RNase H domain in the p51 homodimer may yield a molecule with greater conformational flexibility compared to the heterodimer. The lag time observed before the onset of polymerization observed with p51 suggests that template–primer binding itself may be important for the attainment of the active polymerase conformation for this molecule. In the case of the heterodimer, the interaction of the p51 subunit with the RNase H domain of the p66 subunit would induce or stabilize the locked conformation of p51 subunit in the absence of template–primer.

Results from the primer extension and processivity analyses suggest that heterodimeric RT reinitiates synthesis more rapidly than p51 homodimeric RT. At short reaction times for premixed template–primer and enzyme, both RT forms yield the same ladder and maximum length of extension products, indicating that the modes and rates of synthesis are similar. At later reaction times, p66/51 extends all products to the average maximum size of the template, whereas p51 leaves a distribution of intermediate sized fragments. Although strand transfer has been shown to be a viable mechanism for primer extension (Huber *et al.*, 1989) and could contribute to the differences seen in this study, the reaction conditions used here, for example, 30 °C and 50 mM KCl, appear not to favor this process. On the basis of homologous RT proteins with deleted or deficient RNase H function (Buiser *et al.*, 1991), we do not expect a substantial difference in strand-transfer activity for p51 and p66/51 in any case.

Processivity measurements from template challenge experiments show a 2-fold higher average processivity of the heterodimer over the p51 homodimer. These results could be explained by either a less rapid k_{on} or a more rapid k_{off} of the enzyme–template–primer complex for p51 homodimer versus heterodimer. Since the template challenge experiments (Figure 7) give a k_{off} value approximately 2-fold lower for p51 than for p66/51, the off-rates of the complexes apparently do not account for the difference in the product extension patterns. A change in the enzyme conformation exclusive to the p66/51 protein that leads to higher processivity might have been expected to give a lower k_{off} for the heterodimer. Reardon *et al.* (1991) have shown that the affinity for template–primer of heterodimer is 20 times larger than that for p51 homodimer, so that apparently k_{on} is on the order of 40 times faster for heterodimer than for homodimer. These results point to the involvement of the RNase H domain in stabilizing the conformation of the heterodimeric enzyme for polymerization and are consistent with the substantial contact of the RNase H domain with the p51 domain seen in the X-ray structure.

The strong primer length dependence displayed by the heterodimer and not by the p51 homodimer (Reardon *et al.*, 1991) suggests that, alternatively, the RNase H domain may stabilize p66/51 *via* its interactions with the template–primer. However, in this case, the higher k_{off} for the heterodimer is more difficult to reconcile.

In conclusion, we have expressed and purified highly active p51 homodimeric RT. Comparison of the enzymatic properties of p51 and p66/51 RT indicates that this homodimer has a polymerase activity comparable to that of the heterodimer and that their structural conformations are therefore likely to be highly similar. The principal effect of the RNase H domain on polymerase function is that it contributes to the stabilization of a p51 subunit conformation in the enzyme necessary for rapid formation of the Michaelis complex.

ACKNOWLEDGMENT

We thank A. Lindgren-Berndt for technical assistance.

REFERENCES

Anderson, S. F., & Coleman, J. E. (1992) *Biochemistry 31*, 8221–8228.
Barber, A. M., Hizi, A., Maizel, J. R., & Hughes, S. H. (1990) *AIDS Res. Hum. Retroviruses 6*, 1061–1072.
Basu,A., Basu, S., & Modak , M. J. (1992) *Biochem. Biophys. Res. Commun. 183*, 1131–1138.
Bavand, M. R., Feitelson, M., & Laub, O. (1989) *J. Virol. 63*, 1019–1021.
Becerra, S. P., Kumar, A., Lewis, M. C.,Widen, S. G., Abbotts, J., Karawya E. M., Hughes, S. H., Shiloach, J., & Wilson, S. H. (1991) *Biochemistry 30*, 11707–11719.
Bradford, M. M. (1976) *Anal. Biochem. 72*, 248–254.
Bryant, F. R., Johnson, K. A., & Benkovic, S. J. (1983) *Biochemistry 22*, 3537–3546.
Buiser, R. G., DeStefano, J. J., Mallaber, L. M., Fay, P. J., & Bambara, R. A. (1991) *J. Biol. Chem. 266*, 13103–13109.
Chen, N., Painter, G. R., & Furman, P. A. (1991) *Biochem. Biophys. Res. Commun. 174*, 785–789.
De Stefano, J. J., Buiser, R. G., Mallaber, L. M., Myers, T. W., Bambara, R. A., & Fay, P. J. (1991) *J.Biol. Chem. 266*, 7423–7431.
Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.
De Vico, A. L., Copeland, T. D., Oroszlan, S., Gallo, R. C., & Sarngadharan, M. G. (1991) *J. Biol. Chem. 266*, 6774–6779.
DiMarzo Veronese, F., Copeland, T. D., De Vico, A. L. Rahman, R., Oroszlan, S., Gallo, R. C., & Sarngadharan, M. G. (1986) *Science 231*, 1289–1291.
Dirani-Diab, R. E., Andreola, M. L., Nevinski, G., Tharaud, D., Barr, P. J., Litvak, S., & Tarrago-Litvak, L. (1992) *FEBS Lett. 301*, 23–28.
Dudding, L. R., Nkabinde, N. C., & Mizrahi, V. (1991) *Biochemistry 30*, 10498–10506.
Eisenberg, H., & Felsenfeld, G. (1967) *J. Mol. Biol. 30*, 17–25.
Furman, P. A., Painter, G., Wilson, J. E., Cheng, N., & Hopkins, S. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 6013–6017.
Gill, S. C., & v.Hippel, P. H. (1989) *Anal. Biochem. 82*, 319–326.
Graves, M. C., Meidel, M. C., Pau,Y.-C. E., Manneberg, M., Lahan, H.-W., & Gruninger-Leich, F. (1990) *Biochem. Biophys. Res. Commun. 168*, 30–37.
Hansen, J., Schulze,T., Mellert, W., & Moelling, K. (1988) *EMBO J. 7*, 239–243.
Hizi, A., & Shaharabany, M. (1990) *J. Biol. Chem. 267*, 18255–18258 .
Hizi, A., McGill, C., & Hughes, S. H. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1218–1222.
Hostomsky, Z., Hostomska, Z., Fi, T.-B., & Taylor, J. (1992) *J. Virol. 66*, 3179–3182
Howard, K. J., Frank, K. B., Sim, I. S., & LeGrice, S. F. (1991) *J. Biol. Chem. 266*, 23003–23009.
Hu, Y.-W., & Kang, Y. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4596–6000.
Huber, H. E., McCoy, J. M., Seehra, J. S., & Richardson, C. C. (1989) *J. Biol. Chem. 264*, 4669–4678.
Jacobo-Molina, A., & Arnold, E. (1991) *Biochemistry 30*, 6351–6361.
Kedar, P. S., Abbott, S. J., Tovacs, T., Lesiak, K., Torence, P., & Wilson, S. H. (1991) *Biochemistry 29*, 3603–3611.
Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., & Steitz, T. A. (1992) *Science 256*, 1783–1790.
Krug, M. S., & Berger, S. L.(1991) *Biochemistry 30*, 10614–10623.
LeGrice, S. F. J., Naas, T., Wohlgesinger, B., & Schatz, O. (1991) *EMBO J. 10*, 3905–3911.

Lori, F., DiMarzo Veronese, F., DeVico, A. L. Lusso, P., Reitz, M. S., & Gallo, R. C. (1992) *J. Virol. 66*, 5067–5074.
Majumdar, A., Abbotts, J., Broder, S., & Wilson, S. H. (1988) *J. Biol. Chem. 263*, 15657–15665.
Mueller, B., Restle, T., Weiss, S., Gautel, M., Sczakiel, G., & Goody, R. S. (1989) *J. Biol. Chem. 264*, 13975–13978.
Panganiban, A. T., & Fiore, D. (1988) *Science 241*, 1064–1069.
Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, J. S. R., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C., & Wong-Staal, F. (1985) *Nature 313*, 277–284.
Reardon, J. E. (1992) *Biochemistry 31*, 4473–4479.
Reardon, J. E., & Miller, W. H. (1990) *J. Biol. Chem. 265*, 20302–20307.
Reardon, J. E., Furfine, E. S., & Cheng, N. (1991) *J. Biol. Chem. 266*, 14128–14134.
Restle, T., Mueller, B., & Goody, R. S. (1990) *J. Biol. Chem. 265*, 8986–8988.
Saag, M. S., Hahn, B. H., Gibbons, J., Li, Y., & Parks, E. S., Parks, W. P., & Shaw, G. M. (1988) *Nature 334*, 440–447.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular cloning*, 2 ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Schein, C. H., & Noteborn, M. H. M. (1988) *Bio/Technology 6*, 291–294.
Sobol, R. W., Suhadolnik, R. J., Kumar, A., Lee, B. J., Hatfield, D. C., & Wilson, S. H. (1991) *Biochemistry 30*, 10623–10631.
Soutschek, E., Motz, M., Wagner, R., & Wolf, H. (1989) Poster Th.C.T.73, 5th International Conference on AIDS, Montreal, Quebec, Canada.
Starnes, M. C., Gao, W., Ting, R. Y. C., & Cheng, Y.-C. (1988) *J. Biol. Chem. 263*, 5132–5134.
Tisdale, M., Ertl, P., Larder, B. A., Purifoy, D. J. M., Darby, G., & Powell, K. L. (1988) *J. Virol. 62*, 3662–3667.
Weiss, R., Teich, N., Varmus, H., & Coffin, J. (1982) *RNA Tumor Viruses*, p 381, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Wong, I., Patel, S. S., & Johnson, K. A. (1991) *Biochemistry 30*, 526–537.